Inventors: Samuel Naffziger; Donald C. Soltis, Jr.

A METHOD AND SYSTEM FOR DETECTING DROPPED MICRO-PACKETS

BACKGROUND OF THE INVENTION

The present invention relates to data transmissions between agents in a network and computer interconnect fabric.

Transmissions between agents in a typical network or computer inter-connect fabric are done using "packets" which generally comprise two or more flits or micro-packets that are usually rather small, e.g., 128 bits, to ensure a short transmission time and enable easy handling by very large scale integrated (VLSI) chips along the path. In addition to the data, they contain a small control portion which contains information about the destination locations of the flit and perhaps other information. Dropped flits indicate a failure mode that is not detected by standard cyclic redundancy checking (CRC) or error correction codes (ECC) methods. Parenthetically, such dropped flits can be caused by soft errors in VLSI chips that route the flit to the wrong destination or cause it to be ignored by one of the routers. In this context, soft errors refer to stored information that is lost due to high energy particles resulting from radioactive decay (alpha particles) or gamma rays.

Prior art methods of ensuring the reliability of packet transmissions fall into two categories, i.e., flit-level error detection and correction and end to end transmission assurance. Cyclic redundancy check or error correcting codes can check the contents of a flit for errors in transmission, and depending on the code used and the nature of the error, can make corrections. This approach works well to handle error events that operate on the bit level such as electrical noise coupling on the wires used to transmit the data, or random bit flipping in the data portion of the flit.

The end to end transmission assurance involves an acknowledgement sequence between the ultimate recipient of a packet and the sending agent. With this method, the receiver of a packet immediately sends an acknowledgement packet to the sender when the complete packet is received. The sending agent must hold a complete copy of each packet sent until the acknowledgement packet is received. This approach works well in handling a large class of errors that can corrupt a packet during its transmission. The cost, however, is high since the sending agent must store all packets that are in flight and must use some sort of time out mechanism to determine if the receiver has not gotten the packet, at which time the sender is required to resend the packet. In addition there is the overhead of the acknowledgement packets consuming extra bandwidth in the network.

A need exists to easily detect dropped flits.

SUMMARY OF THE INVENTION

The present invention comprises a system and a method of providing error detection and correction of transmission of multiple flits between sending and receiving agents connected together in a network or computer interconnect environment that comprises embedding a sequence identifier in each flit prior to

transmission, sending each flit to a connected receiving agent, examining the sequence identifiers of each flit being received and requesting the sending agent to resend a flit if the sequence identifier for that flit is determined to be incorrect.

In a preferred embodiment of the present invention, the sequence identifier is embedded in the control portion of the flit and comprises a sequence number that is incremented or otherwise changed in a predictable manner, so that the order of flits being received is predicted. If the sequence number for a flit is different than expected, the receiving agent requests that it be resent.

DESCRIPTION OF THE DRAWINGS

FIGURE 1 is a diagram of a data packet comprising a multiplicity of flits having a control portion and a data portion.

FIG. 2 is a diagram of an example of a network with dual processor nodes, and particularly illustrating a packet transmission utilizing multiple hops between two nodes.

DETAILED DESCRIPTION

The present invention comprises a complimentary error detection and correction approach to the prior art methods of end to end transmission assurance and flit-level error detection and correction, such as cyclic redundancy check and error correcting codes. It is believed to provide a lower cost solution than end to end transmission assurance, but more robust method than flit-level error detection and correction. Failure modes that would not be caught by the flit-level error detection and correction method include errors in VLSI circuitry or wires causing corruption of the control portion of the flit and errors in VLSI circuitry causing the flit to be dropped or lost in its entirety.

The system and method of the present invention is intended for use in the transmission of packets comprised of multiple flits that are transmitted over one or more hops, i.e., crossing one or more agents, to arrive at a destination agent. In this regard, an agent is a processor or other VLSI chip such as a memory controller or input/output (I/O) controller connected in a multiprocessing network or fabric. As shown in FIG. 1, which diagrammatically illustrates a network with dual processor nodes and particularly illustrates transmission of from agent 10 to agent 12, a flit must traverse hops between agents 14, 16, 18 and 20. In the drawing, agents 16 and 18 are directly connected together.

In the present invention, and referring to FIG. 2, a packet 22 typically comprises a plurality of flits that may number from 2 to N with each flit having a control portion 24 and a data portion 26. The control portion 24 may have several fields of information such as origination information, destination and other information (not shown), but importantly to this invention a sequence identifier that is changed in a predictable manner so that the order in which flits are sent and received can be determined. While the sequence identifier may be changed in any predictable manner, the preferred embodiment merely increments a number by 1 for successive flits. This is carried out by an algorithm which in pseudo-code comprises:

```
if (new flit received {
if (flit==data flit && flit != header flit) {
Extract sequence number -> s_new;
if (s_new != s_old+1 {
{ signal error to sender:
} s_old = s_new;
}
}
```

While the foregoing algorithm is used in the preferred embodiment, any predictable incrementing or decrementing operation or digital signature or computation that enables the order of flits to be determined is within the scope of the present invention.

To detect dropped flits, the present invention in its preferred embodiment embeds a sequence number in each flit, incremented up from a value that is substantially unique for each packet. As each agent along the transmission path from sender to receiver gets the flit, it checks that this sequence number is the next in line for the packet to which it belongs. If an out of order flit is received, the agent receiving it sends a request for resend to the sending agent, which is not necessarily the original sender.

When a sequence number mismatch is detected at the receiving agent, it then signals the sending agent of a failure. This means the sending agent is required to hang on to at least one extra flit in a replay buffer to be able to resend the dropped flit since an error isn't detected until after the subsequent flit is sent. In this regard, whether a copy of the flit is written into a separate replay buffer or merely retained in a memory location is largely a matter of semantics in that one of ordinary skill in the art can manipulate the flit to accomplish the retention and resending of the flit and many alternative types of manipulation is within the scope of the invention. Importantly, the amount of storage required in each agent is quite small since the re-send operation is at an agent-to-agent level, not sender to receiver. In addition, a time out mechanism is avoided since every hop on the transmission path requires either an acknowledgement, or error indication. Such communication can be arranged to consume only a single wire since it is between connected agents in the network.

Another benefit of the present invention is that a catastrophic failure of a VLSI chip somewhere in the transmission path will be detected as a missing

or incomplete sequence number. This will allow the destination agent to recognize that an error has occurred in this packet and flag the error instead of continuing to consume information with silently corrupted data.

From the foregoing, it should be appreciated that a system and method of providing error detection and correction of transmission of multiple flits between sending and receiving agents has been described that has many desirable attributes and advantages compared to known prior art systems. The present invention provides a low cost solution for reliably detecting and correcting errors in transmission of flits that are incapable of being detected and corrected by known prior techniques.

While various embodiments of the present invention have been shown and described, it should be understood that other modifications, substitutions and alternatives are apparent to one of ordinary skill in the art. Such modifications, substitutions and alternatives can be made without departing from the spirit and scope of the invention, which should be determined from the appended claims.

Various features of the invention are set forth in the following claims.